



AIMSweb® Training Workbook:

Organizing and Implementing a Benchmark Assessment Program

Mark R. Shinn, Ph.D.

Edformation, Inc.
6420 Flying Cloud Drive,
Suite 204
Eden Prairie, MN 55344

P: (952) 944-1882 (888) 944-1882 **F:** (952) 944-1884

www.aimsweb.com

©2002 Edformation Inc. All rights reserved. No part of this book may be reproduced or transmitted in any form by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher.

Table of Contents

This manual is to be used as a supplement to the *AIMSweb Power Point Training Presentation* and *AIMSweb Training Video*.

Letter to AIMSweb Subscriber	3
Overview of AIMSweb Training Materials	5
Big Ideas About General Outcomes Measurement (GOM)	6
Summative Evaluation - Inform the Public	7
Formative Evaluation - Inform Instruction	8
AIMSweb Benchmark Assessment: Making Formative Evaluation Feasible	9
AIMSweb Benchmark Assessment: Providing More than Formative Evaluation	14
AIMSweb Benchmark Assessment: Useful for Communicating with Parents	15
AIMSweb Benchmark Assessment: Useful for Principals and Administrators	16
An Extended Example: Teacher Benchmark Reports	18
Formative Evaluation: Inform Instruction	22
Setting It All Up	22
When to Test	23
What to Test and Score	24
Who to Test	25
Who Does the Testing	25
Professional Ethics: Test Security, Confidentiality, and Sensitivity	27
Training	28
How Students are Tested	29
After Testing: What Then?	34
Summary	34
References	35
Appendix	36
Letter to Parents	36
Reading Benchmark Testing Schedule	39
Benchmark Assessment Essentials Checklist	40

Dear AIMSweb Subscriber:

Welcome to the *AIMSweb* formative assessment and basic skills improvement system. *AIMSweb* provides teachers, school administrators, and parents a complement to the summative (high stakes) assessment/evaluation model prevalent in education today. Rather than just providing schools with information about student learning at the end of the school year, *AIMSweb* organizes and reports the results of simple, accurate, low cost, and more frequent testing using validated General Outcome Measures like Curriculum-Based Measurement during the school year. The *AIMSweb* formative assessment model informs the instructional process as it occurs by identifying at risk students as early as possible and importantly, those students who are learning and those who are not progressing satisfactorily. The distinction between “did they learn last year” and “are they learning this year” represents a paradigm shift, one that is critical for quality improvement!

The *AIMSweb* system consists of four components:

1. Two web-based data management and information reporting programs to report and graph the results of Curriculum-Based Measurement (CBM) in early literacy, reading, and spelling.
 - *AIMSweb Benchmark* manages, evaluates, reports, and charts the results of three times per year school benchmark assessments for all students Grades K-8.
 - *AIMSweb Progress Monitor* allows teachers to monitor students at risk or those students with more severe educational needs more frequently to evaluate the effects of interventions and document appropriate instructional changes.
2. Standard General Curriculum Assessment Materials:
 - *Standard Benchmark Reading Assessment Passages:* A set of 3 graded and equivalent standard passages for Grades 1-8 for establishing fall, winter and spring reading benchmarks (24 total passages). These passages are also available in Spanish.
 - *Standard Progress Monitoring Reading Assessment Passages:* A set of 30 graded and equivalent passages for Grades 2-8, 20 for Grade 1 and 20 for Primer Level for use in more frequent and continuous monitoring (250 passages total).
 - *Early Literacy Measures:* A set of 3 equivalent Standard Benchmark Early Literacy Indicators to assess Phonemic Awareness and Phonics for Kindergarten and Grade 1 for establishing fall, winter, and spring benchmarks.
 - *Early Literacy Indicators for Progress Monitoring:* A set of 20 equivalent Standard Early Literacy Indicators for Kindergarten and Grade 1 for use in more frequent and continuous monitoring of early literacy skills (20 tests for each indicator).
 - *Standard Maze Passages:* Three Standard Assessment Reading Passages for Grades 1-8 have been prepared in a maze (multiple choice close) format for use as another measure of reading comprehension (24 maze passages total).

- ***Standard Benchmark Spelling Lists:*** A set of graded and equivalent standard spelling lists for use in Grades 1-8 for establishing fall, winter, and spring spelling benchmarks (24 total lists).
- ***Standard Progress Monitoring Spelling Lists:*** A set of 30 graded and equivalent lists of Grade 2-8 spelling words per grade and 20 lists of Grade 1 words (230 total) for use in more frequent and continuous monitoring.

3. Training Workbooks designed to train staff to implement the *AIMSweb* system.

- ***Administration and Scoring of Reading Curriculum-Based Measurement (R-CBM) for Use in General Outcome Measurement***
- ***Administration and Scoring of Early Literacy Measures for Use in General Outcome Measurement***
- ***Administration and Scoring of Spelling Curriculum-Based Measurement (S-CBM) for Use in General Outcome Measurement***
- ***Administration and Scoring of Reading Maze for Use in General Outcome Measurement of Reading Comprehension***
- ***Organizing and Implementing a Benchmark Assessment Program***
- ***AIMSweb Progress Monitor - Strategies for Writing Individualized Goals in General Curriculum and More Frequent Formative Evaluation***

AIMSweb trainers are available to deliver the training onsite or the materials can be used without assistance.

4. Online Support:

AIMSweb users become members of a community of users and an online support site (*AIMSonline*) designed to solve problems, answer questions, and contribute to professional development and successful implementation. A network of Strategic School Partners and Certified AIMSweb Trainers located around the country are available for inquiries, expertise, training, onsite visits, etc. *AIMSweb* "informs" the teaching and learning process by providing continuous student performance data and reports improvement to students, parents, teachers, and administrators.

Our promise to you is simple. Use of the *AIMSweb* system will improve instruction, increase achievement, and report improvement to all stakeholders.

Gary Germann
President/CEO

Steve Jennen,
Vice President and Chief Technical Officer

Overview of AIMSweb Training Materials

This is one in a series of Training Workbooks developed to accompany *AIMSweb* (Achievement Improvement Monitoring System). The purpose of the series is to provide instruction, delivery models, and practice opportunities to better use *AIMSweb* to improve achievement outcomes.

Administering and Scoring of Reading Curriculum-Based Measurement (R-CBM) for Use in General Outcome Measurement provides instruction and practice in the skill area of reading. The workbook is accompanied by the AIMSweb Practice Video which contains segments of students reading to demonstrate key features of administering and scoring the graded reading tests. Critical activities to complete before, during, and after testing, including scoring rules, are provided. Practice examples and answer keys are provided for users to observe and score as well as reproducible forms for making testing easier and more accurate. A Power Point Presentation accompanies the user through the training experience.

Administering and Scoring of Spelling Curriculum-Based Measurement (S-CBM) for Use in General Outcome Measurement provides instruction and practice in the skill area of spelling. The workbook is to be used with the AIMSweb Practice Video which also contains demonstrations of key features of administering the graded spelling lists. Critical activities to complete before, during, and after testing, including scoring rules, are provided. Practice examples and answer keys are provided for users to observe and score as well as reproducible forms for making testing easier and more accurate. A Power Point Presentation accompanies the user through the training experience.

Administering and Scoring of Early Literacy Measures for Use in General Outcome Measurement provides instruction and practice in the skill areas of early reading. The workbook describes five fluency measures designed to assess early literacy acquisition from early Kindergarten to Grade 1, including Beginning Sounds, Letter Names, Letter Sounds, Phonemic Segmentation, and Nonsense Words. The workbook is accompanied by a videotape of students taking these tests to demonstrate key features of administering and scoring each indicator. Critical activities to complete before, during, and after testing, including scoring rules, are provided. Practice examples and answer keys are provided for users to observe and score as well as reproducible forms for making testing easier and more accurate. A Power Point Presentation accompanies the user through the training experience.

Administering and Scoring of Reading Maze for Use in General Outcome Measurement provides instruction and practice in the skill area of reading comprehension. Critical activities to complete before, during, and after testing, including scoring rules, are provided. Practice examples and answer keys are provided for users to observe and score as well as reproducible forms for making testing easier and more accurate. A Power Point Presentation accompanies the user through the training experience.

Organizing and Implementing a Benchmark Assessment Program provides information on how to conduct benchmark testing in general education classrooms. The workbook provides straightforward, simple, and valuable information for planning, communication, and conducting all school benchmark testing. This manual is intended for use with *AIMSweb Benchmark* web-based software.

AIMSweb Progress Monitor - Strategies for Writing Individualized Goals in General Curriculum and More Frequent Formative Evaluation instructs teachers on how to write individualized annual goals for students and monitor progress on a frequent and continuous basis. Intended for use with students in individualized remedial programs - such as special education or Title I - the Training Workbook demonstrates how to write individualized annual goals based on a Survey-Level Assessment (SLA) and provides strategies for col-

lecting student outcome information frequently and continuously. This manual is intended for use with the *AIMSweb Progress Monitor* web-based software.

Big Ideas About General Outcome Measurement (GOM)

Medicine measures height, weight, temperature, and blood pressure; the Federal Reserve Board measures the Consumer Price Index; Wall Street measures the Dow-Jones Industrial Average; companies report earnings per share; and even McDonald's measures how many hamburgers they sell. What do these measures have in common? They all assess general outcomes so decisions are data-based and timely.

Although these measures do not assess *all* health, economic, stock market, business or even fast food sales behavior, they are indicators considered so important to outcomes that *they are routine*. These measures are simple, accurate, and reasonably inexpensive in terms of time and materials. They are collected on an ongoing basis over time. They shape a variety of important decisions.

Education has its own set of indicators of general basic skill success. Derived out of the research base generated by a set of test procedures called Curriculum-Based Measurement (CBM), these General Outcome Measures (GOMs) allow us to make important statements about our students' reading, spelling, written expression, and mathematics computation skills.

CBM was developed more than 20 years ago by Stanley Deno at the University of Minnesota, and implemented into schools by Gary Germann, with the idea of giving educators simple, accurate, and efficient indicators of student achievement. School-based research on CBM with real students and real teachers continues to this day. The references included in this workbook provide extensive information about how CBM was developed and validated, and how CBM can be used to make a variety of general and special education decisions.

Originally, CBM was designed to assess growth and development in students' specific curricula. Teachers created their own individual set of CBM passages based on what they were teaching and used the information to determine students' rates of progress and make changes in instruction as needed. This tie to curriculum had high instructional validity but lacked the necessary other technical features of reliable and valid measurement.

It soon became apparent that the positive effects of testing from materials selected from an individual teacher's curriculum were offset by the lack of standard information about students' progress. Some teachers had "no curriculum," the curriculum would change year to year, and the differences between schools, between teachers within schools, and so on, made accurate decisions about students' progress very difficult. Furthermore, teachers were too often burdened by the business of creating their own testing materials. In addition to being more time consuming, the variability in test practices was a concern.

After considerable research, it has been demonstrated that the perfect correspondence between what CBM assessed and students' specific curricula was not necessary. In fact, by using standard test materials the same judgments about students' level of reading skill and reading progress, could still be made accurately, as well as provide appropriate, standards of growth and development across varied curricula, teachers, schools, and school districts.

What emerged from this school-based research was the following conclusion: Achievement can be improved by testing students (1) using standard, valid assessments, (2) that measured something important, (3) on tasks of about equal difficulty tied to general curriculum (4) over time.

CBM provided the testing procedures to be able to do Numbers 1, 2, and 4. By developing graded and equivalent testing materials of about equal difficulty tied to general curriculum, (Number 3) General Outcome Measurements (GOMs) evolved. Thus, the testing procedures known as CBM are used in an testing approach called General Outcome Measurement.

Summative Evaluation - Inform the Public

Improving the achievement of American school children has been a visible, high-priority goal at the local, state, and national level for more than a decade. A number of approaches have been undertaken, including “outcome-based education” and now “high standards-driven” education. All but 1 state have identified their high standards and all states are engaging in high-stakes testing to ensure that students and schools are held to those standards. At set time points, students are assessed and those who pass move to the next grade or graduate. Students who do not pass face a variety of consequences, including retention and/or summer school. Schools are held accountable for passing rates on the high-stakes tests. Too few students passing may result in a variety of sanctions, including teacher or principal transfers, development of school improvement plans, or even loss of accreditation.

High-stakes testing is based on a summative, reactive evaluation approach. Testing occurs after instruction. Summative evaluation is best used for accountability; it provides motivation for students and teachers alike to do their best work. This testing approach is reactive in that it usually takes place well into many students’ academic “careers” (e.g., third grade) when patterns of failure are well established and more difficult to change. It is designed to answer the question “Did students learn?”

Unfortunately, high stakes tests are not useful for making the kinds of instructional decisions teachers need to make every day, such as

- For whom do I need to individualize instruction or find more intensive instructional programs?
- How do I organize my classrooms for instructional grouping?
- Most importantly, how do I know that my teaching is “working” for each student so that I can make changes in instruction when necessary?

High stakes tests also fail to give the information teachers need to:

- Set measurable goals for student progress in the short term (months) vs the long term (years).
- Communicate educational need and progress clearly in non-technical language with parents, and other professionals.
- Determine instructional needs of new students as they constantly appear at school doors.
- Provide information on educational needs, goals, and progress for those students who may need remedial programs like special education.

High-stakes tests also may not be useful to administrators who must make decisions about allocating instructional resources. Testing takes place after years of instruction, with long time frames in between, and well after the information about success and failure rates could be used to make changes in curriculum and instruction. Without the ability to improve the power of education, the primary use of high stakes tests, then, may be to “assign the blame” to students, their parents, teachers, or schools. Accompanying the inability of reactive high stakes tests to inform instruction/teaching is their high cost in terms of loss of instructional time as well as the tests themselves. ***Reactive, summative, high stakes testing may be too little, too late, at too high a cost.***

Formative Evaluation - Inform Instruction

Formative evaluation is the process of assessing student achievement *during* instruction for the purposes of determining whether an instructional program is effective for individual students. When formative tests show that students are progressing, teachers keep using their instructional programs; when tests show that students are not progressing, they can change their instructional programs in meaningful ways. It is designed to answer the question “Are students learning?”

A good way to think about the differences between summative and formative evaluation is with the metaphor of weight. Let’s illustrate summative evaluation. We could set as a high standard that all children have a healthy weight by the end of third grade. We would determine attainment of this standard by a high stakes “test” that precisely calculates each student’s body density. This process would entail weighing each student, immersing each student in a large tub filled with water, and measuring the amount of water displaced. Weight would be divided by displacement and this statistic would give us a very accurate picture of physical status. After approximately 8-9 years of growing, we would test the children, place students who are “unhealthy” in remedial programs, create new health programs, blame the ineffectiveness of old health programs or blame these students (or their families) for over- or under-eating.

Alternately, from a formative evaluation perspective we could monitor students’ weight from a Well Check perspective; that is, from birth we would measure a child’s weight continuously with a simple, albeit less precise, general outcome measure - the weight scale. At any point in development, the child could be weighed and a decision about healthy weight could be made. This process is efficient, accurate, and proactive. We would know along the way whether the child is low or high weight and we could change diets, eating habits, and/or exercise programs long before the high stakes point.

The use of formative evaluation systems in education has been linked to important gains in student achievement (L. Fuchs, 1986) with effect sizes of .7 and greater. An effect size of .7 would mean that a student who was at the 50th percentile without formative evaluation would be expected to perform at the 76th percentile with formative evaluation. This impact on learning is impressive.

Formative evaluation requires the use of repeated tests over time. These tests cannot be just “any tests” but need to have some critical features such as those described in a recent professional article by Fuchs and Fuchs (1999):

- Technical adequacy (reliability and validity);
- Capacity to model growth (able to represent student achievement growth within and across academic years);
- Treatment sensitivity (scores should change when students are learning);
- Independence from specific instructional techniques (instructionally eclectic so the system can be used with any type of instruction or curriculum);
- Capacity to inform teaching (should provide information to help teachers improve instruction);
- Feasibility (must be doable).

The General Outcome Measures, Curriculum-Based Measurement (CBM) used in the ***AIMSweb System*** were designed specifically to match these characteristics and to be used in formative evaluation. More information on research evidence regarding CBM can be found in the journal articles and book chapters listed in the Reference Section.

AIMSweb Benchmark Assessment: Making Formative Evaluation Feasible

The use of General Outcome Measurement in formative evaluation has its roots in special education where there is both (a) legal demands for frequent progress monitoring toward individualized annual IEP goals and reporting to parents, and (b) a body of knowledge showing that formative evaluation strategy improves student achievement outcomes with students with disabilities. In reading, students are tested using Edformation's **Standard Progress Monitor Reading Assessment Passages (RAPs)** once or twice per week. This process is the core of **AIMSweb Progress Monitor**.

Feasibility is less of a concern in remedial programs because class sizes are smaller and there are more personnel resources, including trained aides, than in general education. We increase feasibility in general education by collecting CBM information in Benchmark Assessment. Benchmark Assessment is a process where information is collected on all students in general education classrooms, including students in remedial programs, at regular intervals, usually during the fall, winter, and spring of an academic year with about equal numbers of weeks between assessments. This frequency is less often than in remedial programs but still sufficient to make formative evaluation decisions about students.

In reading, for example, Benchmark Assessment consists of each student reading aloud 3 passages carefully selected for this purpose. For example, all third grade students are assessed using the Grade 3 **Standard Benchmark Reading Assessment Passages**. Each student is also assessed in the Fall, Winter and Spring of each school year. Edformation's **Standard Benchmark Reading Assessment Passages** are designed to correspond to appropriate grade materials and to be equivalent in difficulty.

As will be discussed in more detail later in this workbook, this Benchmark reading information can be collected by a variety of trained persons, including aides, paraprofessionals, school administrators, other professionals, and community volunteers, in a process that takes about 5 minutes total for each student for each Benchmark Assessment. Thus, for each student, Benchmark Assessment in reading requires a commitment of 15 minutes per year to ensure monitoring of reading growth and improvement and instructional success.

Examples of **AIMSweb Benchmark** reports in reading with third graders are presented in Figures 1, 2, and 3. In Figure 1 R-CBM assessment results are shown for Fall, Winter, and Spring Benchmarks for all Grade 3 students in the school. On the figure, the median (50th percentile or the score at which half the students scored higher and half lower) is shown with a black horizontal line. The range of average reading scores (between the 25th and 75th percentile) is shown in green. Scores above average (75th to 90th) are shown with the thick vertical blue line (or top whisker) and scores below average (10th to 25th) are shown with the red bottom whisker. Scores above the blue whisker represent exceptionally high reading scores (greater than the 90th percentile) and scores below the red whisker represent exceptionally low reading scores (less than the 10th percentile).

Note that the reading scores of the third graders as a group in this school are improving Fall (median = 80) to Winter (median = 95) to Spring (median = 120) and this rate of improvement was similar across the range (low average, average, above average) of reading skills.

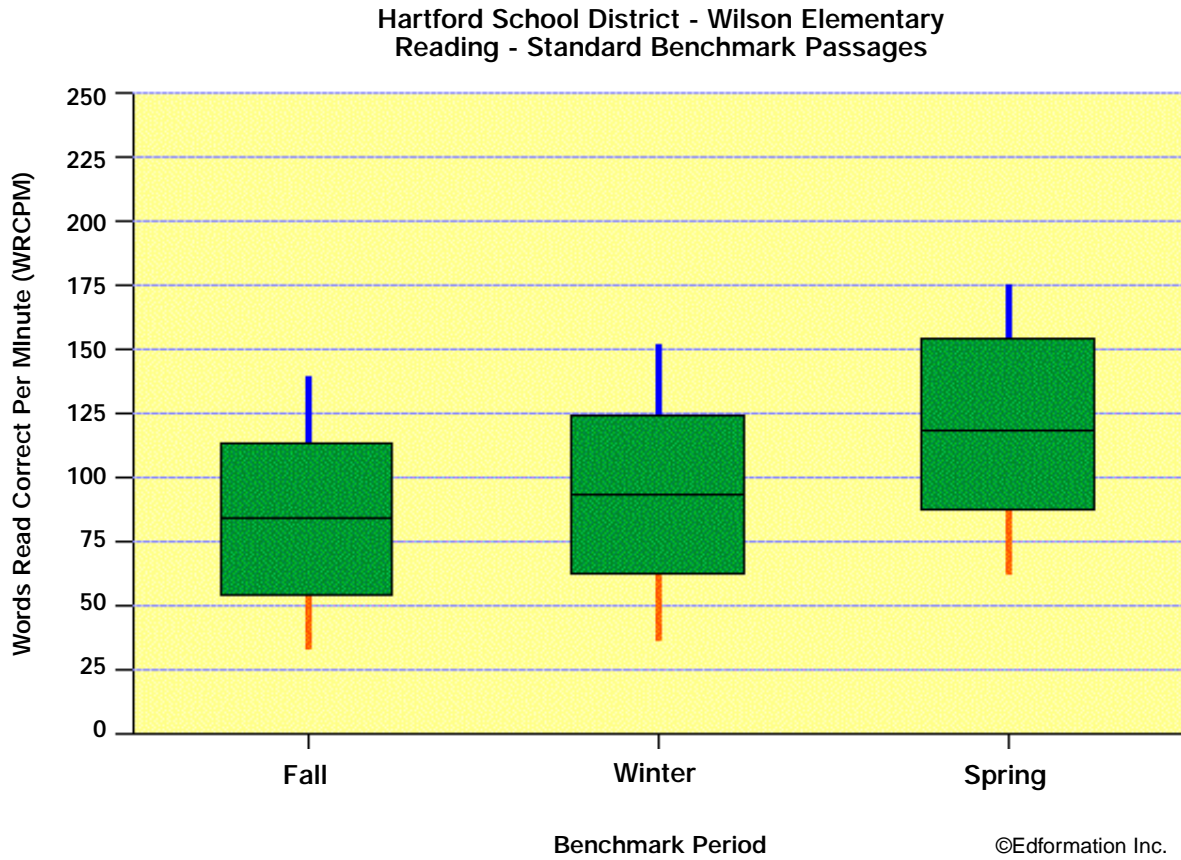


Figure 1

Information on the collective growth of students is interesting, but educators (and parents) want information about the educational needs of individual students.

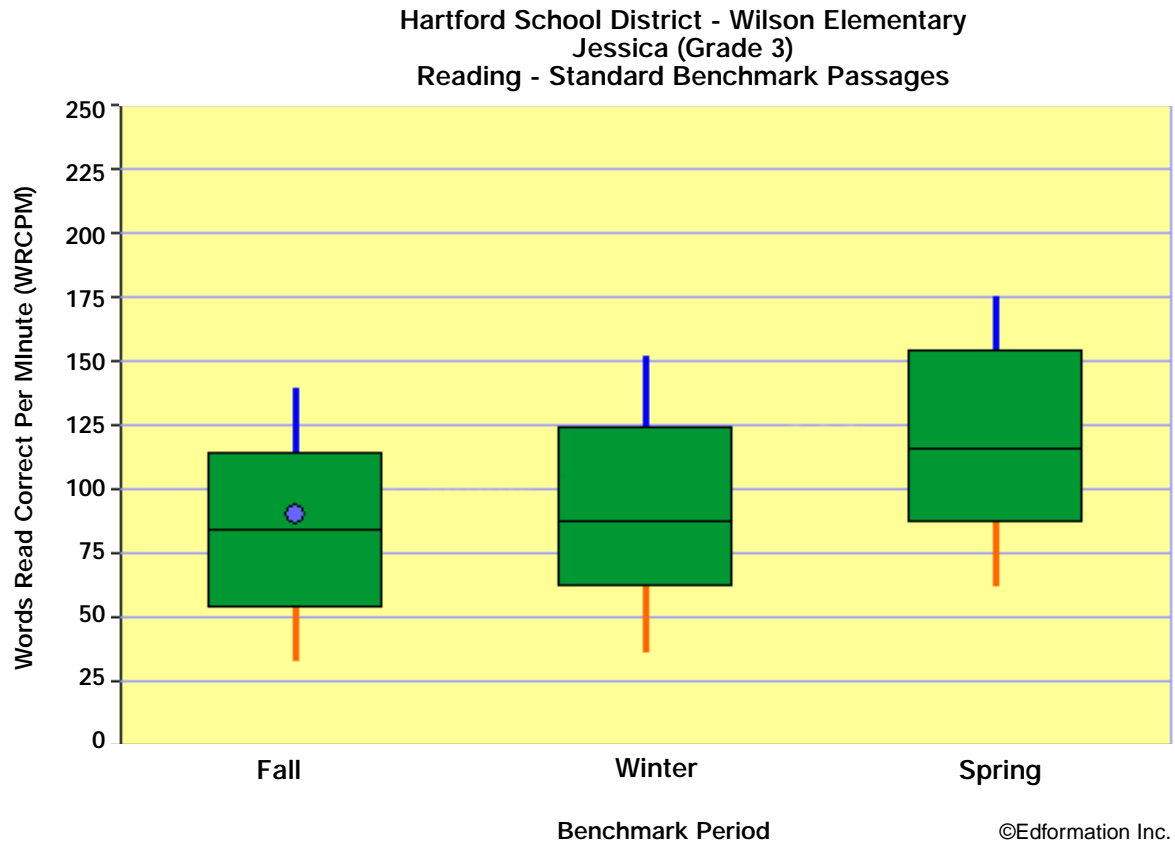


Figure 2

Figure 2 is the *AIMSweb Benchmark Improvement Report* for Jessica. Her scores would be those of an average reader in the fall of Grade 3 in her school.

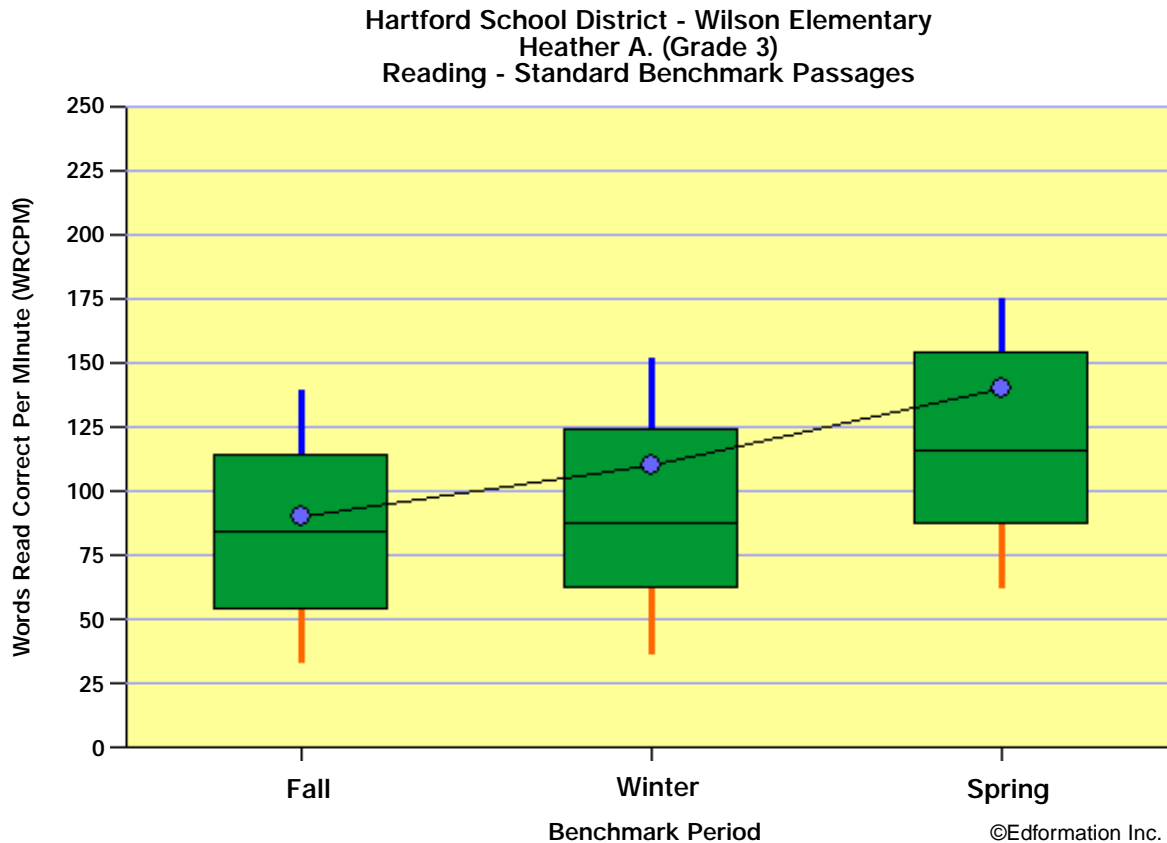


Figure 3

Another example is shown in Figure 3 for Heather.

Figure 3 is the **AIMSweb Benchmark Improvement Report** and shows the assessment results for a specific student, Heather, compared to the Fall, Winter, and Spring Benchmarks for Grade 3 students in her school. In the Fall, Heather read 90 WRC. This score placed her at the 60th percentile. In other words, compared to other students in her school, she was an average reader. At the Winter Benchmark, Heather read 110 WRC. This score placed her at the 66th percentile. Compared to other students in Wilson Elementary School, she was an average reader improving at a slightly faster rate than other students. Given these results, Heather's teacher would keep delivering the same reading program. At the Spring Benchmark, Heather read 140 WRC. This score placed her at the 70th percentile. Compared to other students at Wilson, she again was an average reader improving at the same rate as other students. Given these results, Heather's parents and teacher could conclude with confidence that she benefited from the reading program.

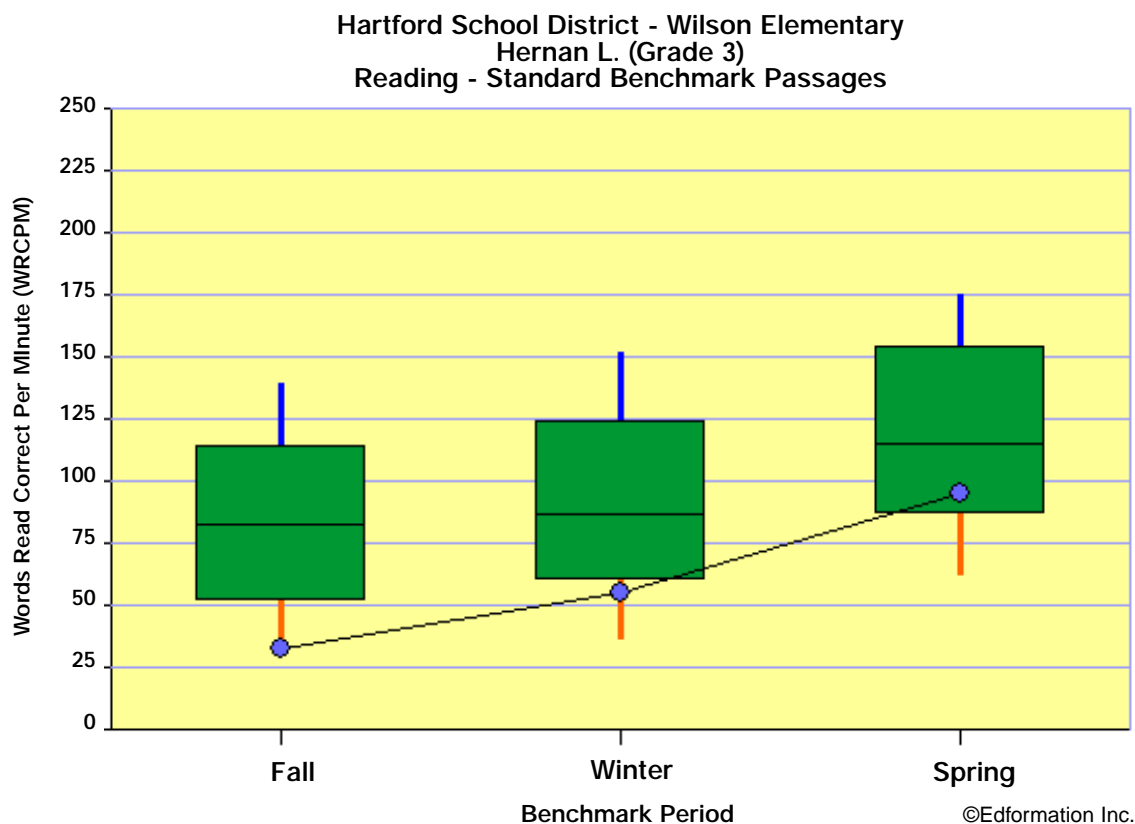


Figure 4

In Figure 4, the **AIMSweb Benchmark Improvement Report** shows Hernan's reading performance compared to the Fall, Winter, and Spring Benchmarks for Grade 3 students at Wilson Elementary School. In the Fall, Hernan read 32 WRC. This score placed him at the 12th percentile. In other words, compared to other students in his school, he was a below average reader and could be considered at risk for lack of reading progress. Because of this potential risk, Hernan's teacher immediately individualized the reading program. At the Winter Benchmark, Hernan read 53 WRC. This score placed him at the 22nd percentile. Compared to other students in Wilson Elementary, he was still an at risk reader, but more importantly, he was improving at a faster rate than other students. Given these results, Hernan's teacher continued the reading program. At the Spring Benchmark, Hernan read 95 WRC. This score placed him at the 30th percentile. Compared to other students in his school, Hernan had improved to a low average reader and was improving at a faster rate than other students. Given these results, Hernan's teacher could conclude with confidence that he benefited from the changes in the reading program and was less at risk.

Benchmark Assessment: Providing More than Formative Evaluation

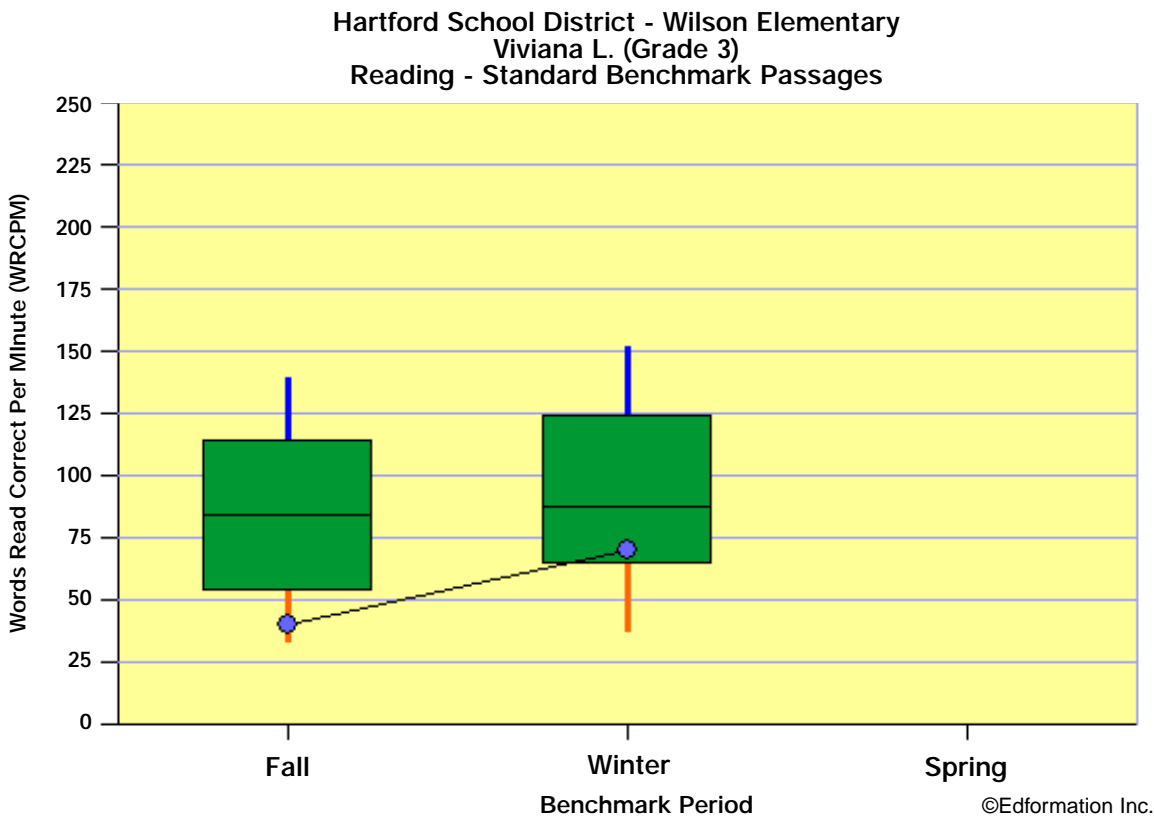
The primary purpose of the **AIMSweb Benchmark System** for teachers is to “inform instruction” through formative evaluation; that is, to let teachers know what is working instructionally and what is not. However, knowing how much each student has improved by the next Benchmark Assessment still is too reactive in many cases. **AIMSweb Benchmark** therefore is also designed to help teachers plan their instruction by making it clear to teachers which students may need more individualized and/or more resource intensive instruction at the beginning of the academic year. We liken this initial Benchmark Assessment to a process of “taking inventory” so teachers can understand the range of skills and the instructional needs of their students as soon as possible. Students whose skills are significantly above or below those of other students may need instructional programs that are different than those provided to more typical students.

Consider, for example, the top reader in a grade. Teachers may want to modify their standard reading program in some meaningful way (e.g., using a more challenging textbook) to allow those students to benefit more fully. Consider as well the lowest readers in that grade. These students could lag so far behind in reading skills that the classroom teacher alone may not meet their instructional needs. In this instance, the classroom teacher may begin problem-solving activities immediately by seeking consultation from a colleague, a building assistance team, or perhaps a reading specialist.

In the case of Hernan, illustrated in Figure 4 previously, he was identified as an at risk reader and his classroom teacher immediately individualized the reading program. How **AIMSweb Benchmark** helps teachers make these kinds of decisions will be described later in the workbook.

Benchmark Assessment: Useful for Communicating with Parents

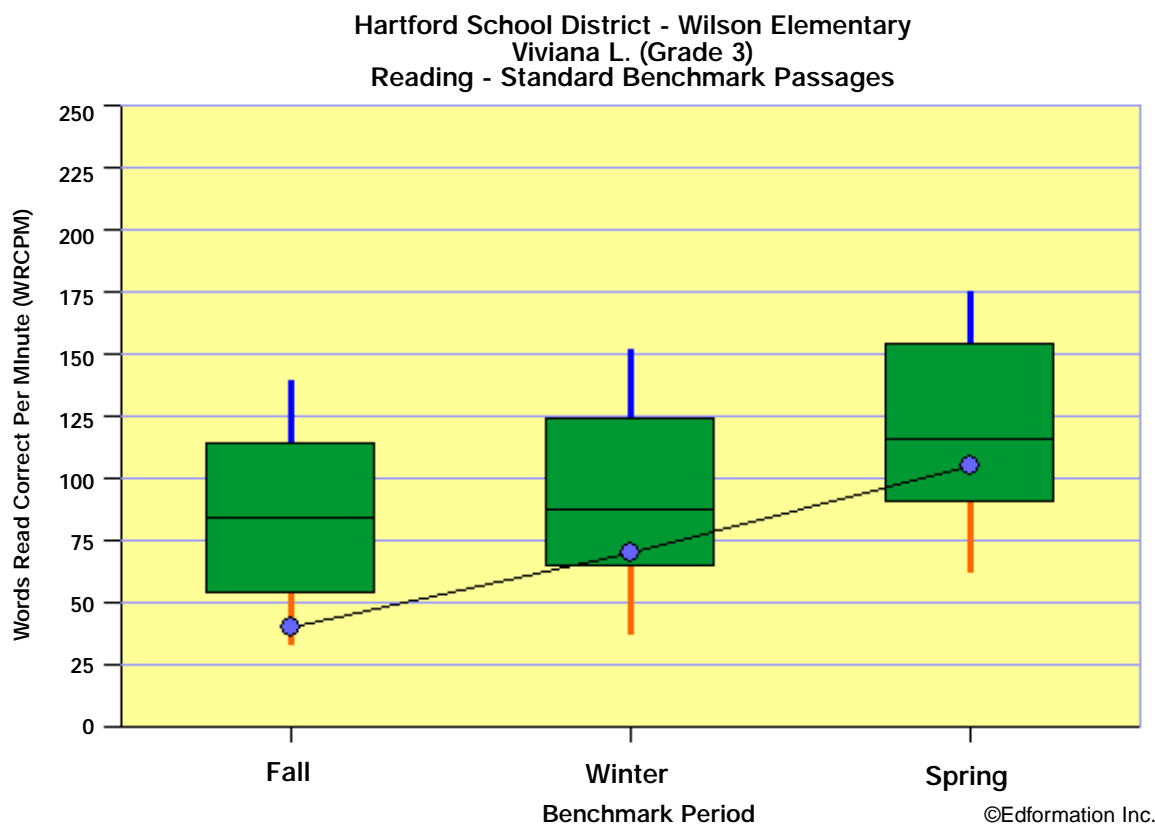
AIMSweb Benchmark also was designed for purposes of collaboration and communication with parents. Based on the assumption that student achievement is enhanced by the degree to which teachers can communicate student achievement growth to parents on a continuous basis, a Benchmark Improvement Report is produced for each Benchmark Assessment. This report, a sample of which comes for the Winter Benchmark in reading is presented in Figure 5 below.



Qualitative Features:	Teacher Comments:
1. Reads very accurately. 2. Reads with expression. 3. Self-corrects errors	Viviana's reading is improving rapidly. She is trying very hard and the results are showing.

Figure 5

Viviana was identified as a at risk reader at the Fall Benchmark, the parents were notified, and the classroom teacher individualized the reading program. The Winter Benchmark Improvement Report shows Viviana's past Fall performance and her current significant improvement. Viviana is now reading in the low average range compared to other third graders in her school and has improved at a faster rate.



Qualitative Features:	Teacher Comments:
<ol style="list-style-type: none"> Has an effective strategy for unknown words. Reads very accurately. Reads with expression. Self-corrects errors 	<p>Viviana has shown excellent improvement in her reading skills this year. Her hard work has paid off. Keep it up and do lots of reading this summer!</p>

Figure 6

The Spring Benchmark Improvement documents continued rapid improvement; Viviana's reading skills have improved so much that she reads like an average reader in her school. The decision to intervene early in the Vivian's reading program resulted in significant, positive reading gains that can be communicated to parents and documented in Viviana's school records.

Benchmark Assessment: Useful for Principals and Administrators

School principals and other administrators are more often asked to make decisions about the "collective good," or groups of students, and about instructional programs such as English Language Learning (ELL) or Title I as a whole than to make instructional decisions for individual students. These decisions typically center on accountability and resource allocation. For accountability, each administrator is responsible for producing quality achievement outcomes. Third-grade students in Wilson Elementary should improve in their reading such that this improvement will allow them to pass with certainty the

high stakes testing in Grade 5. First-grade students in Wilson Elementary should learn to read at a rate equivalent to first-graders in other schools in Hartford School District.

Resource allocation decisions are made about which programs get more, less, or different resources to produce quality achievement outcomes. Title I programs that are failing to show reading improvement may require additional staff development training. ELL programs that produce higher than average rates of reading improvement may be asked to share their instructional materials and methods with programs that are producing less significant reading growth.

Benchmark testing also can allow administrators to look at reading improvement across the grades, within an academic year for purposes of accountability, and between academic years for purposes of continuous improvement. An example of the Administrator's Report showing reading improvement by Benchmark period for Hartford School District is shown in Figure 6. In each grade for each Benchmark, student growth is shown.

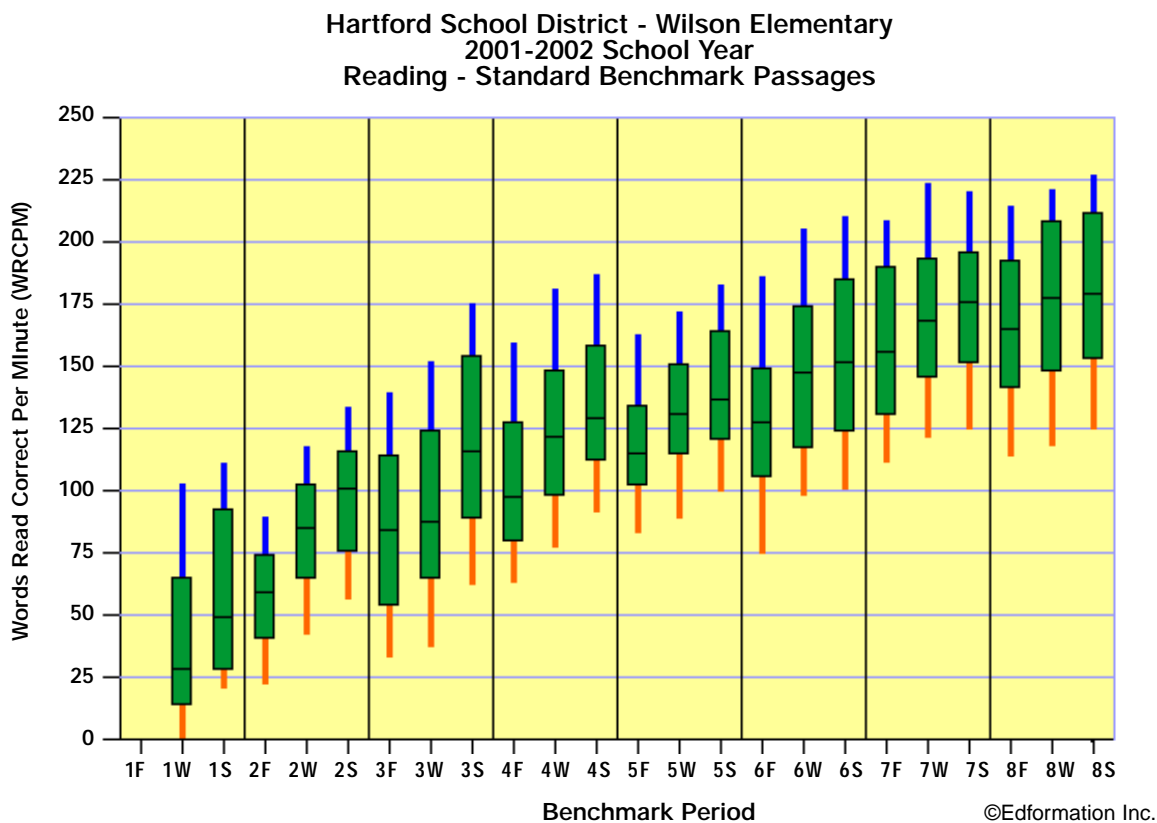


Figure 7

An Extended Example: Benchmark Teacher Reports

To provide a little more understanding of the comprehensive decision-making system that **AIMSweb Benchmark** provides, this section will give the reader a closer look at the Benchmark Teacher Report. This report provides two formats for teachers to see information about students' instructional needs. The first way is in a table format. A sample Benchmark Class Distribution Report from the Fall Benchmark is shown in the following table.

Student Name	Fall WRC	Performance Summary Current Performance	Potential Instructional Action Recommendation
Gamble, K.	184	Above Average Reader	Consider Need for Individualized Instruction
Best, V.	143	Above Average Reader	Consider Need for Individualized Instruction
Lisonbee, I.	142	Above Average Reader	Consider Need for Individualized Instruction
Ruskin, N.	135	Above Average Reader	Consider Need for Individualized Instruction
Quandt, E.	127	Above Average Reader	Consider Need for Individualized Instruction
Dement, B.	126	Above Average Reader	Consider Need for Individualized Instruction
75th%ile = 124			
Damon C.	105	Average Reader	Continue Current Program
Waters, J.	94	Average Reader	Continue Current Program
Smallwood, V.	93	Average Reader	Continue Current Program
25th%ile = 61			
Moon, K.	60	Below Average Reader	Further Assess and Consider Individualizing Program
Mckenney, B.	44	Below Average Reader	Further Assess and Consider Individualizing Program
Ward, D.	43	Below Average Reader	Further Assess and Consider Individualizing Program
10th%ile = 34			
Nankivell, R.	33	Well Below Average Reader	Begin Immediate Problem Solving

This report, based on the Fall reading Benchmark Assessment, ranks the students in each teacher's class from best readers to poorest readers by the number of words read correctly (WRC). The Benchmark Class Distribution Report also color-codes each student by their current reading performance compared to other students in their grade at their school.

The Class Distribution Report provides both a description of each student's current performance and a prescription for Instructional Action(s) that teachers consider for use in instructional decision making. For example, the Fall reading Benchmark Assessment Report describes Damon as an Average Reader in the school. The Fall recommended instructional action for his teacher is to "Continue the Current Reading Program."

Clear	Exceptional Reader	(90th percentile or higher)
Blue	Above Average Reader	(75th to 90th percentile)
Green	Average Reader	(25th to 75th percentile)
Yellow	Below Average Reader	(10th to 25th percentile)
Red	Well Below Average Reader	(Less than the 10th percentile)

The second way that information is presented to teachers is graphically through the Benchmark Improvement Report as shown in Figures 8 and 9. This report is anchored by a graph that shows the performance of each student compared to the others in the grade in the school, as well including the descriptive and prescriptive information in the Class Distribution Report. In the Benchmark Improvement Report, information about what specific instructional actions taken by the teacher is recorded so that changes to meet each student's instructional needs are documented. Damon's Benchmark Improvement Report for the entire school year is shown below. He started the year as an Average Reader, and improved at a faster rate than other students from Fall to Winter, resulting in his status at that time as an Above Average Reader. He progressed less rapidly from the Winter to Spring Benchmark, ending the year with significant progress and maintaining his performance level as an Average Reader compared to other students at Wilson Elementary.

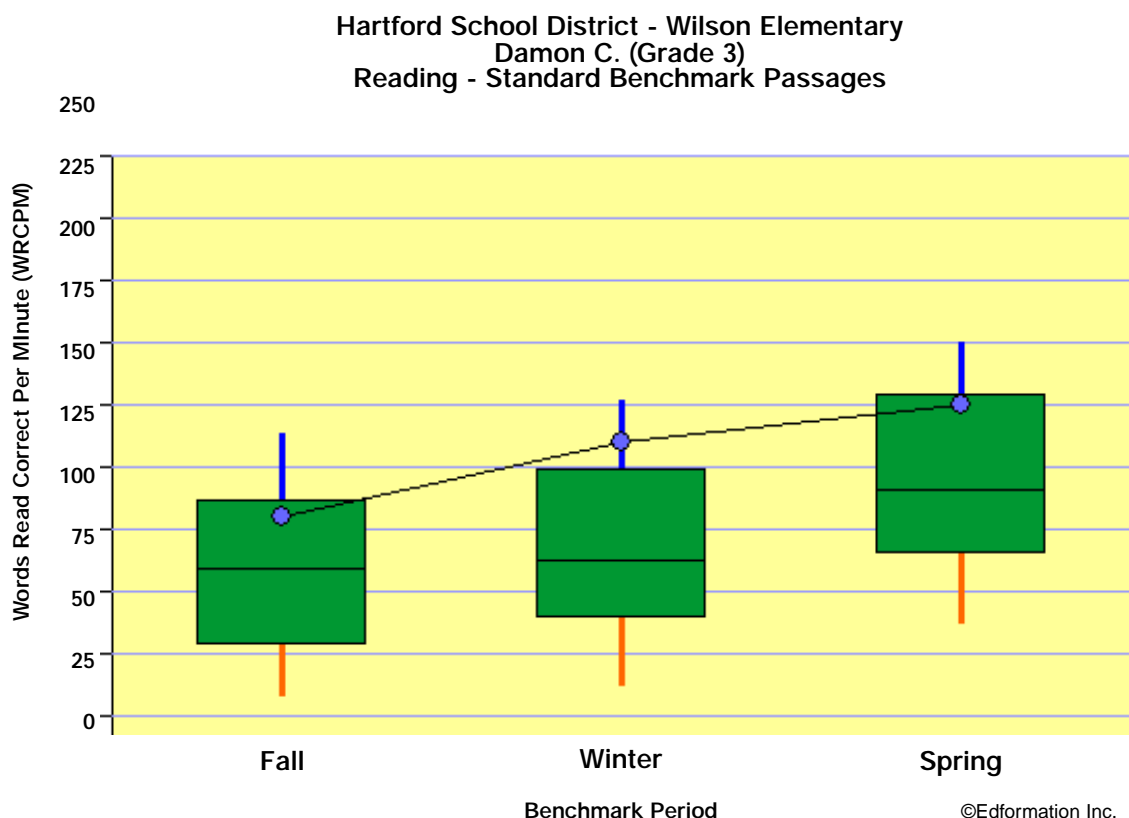


Figure 8

Qualitative Features:	Teacher Comments:
<ol style="list-style-type: none"> 1. Reads fluently or efficiently. 2. Reads with expression. 3. Has an effective strategy for unknown words. 4. Self-corrects errors 5. Adjusts pace when complexity of "considerateness" of text changes. 	<p>Damon improved his reading at the same rate as other students. He needs to continue practicing with multi-syllabic words and common suffixes.</p>

U. Boardman on the other hand, began the year an Above Average Reader with a prescription for the teacher to Consider Need for Individualized Instruction. (Figure 9) His teacher used whole group instruction that he felt was suitable for the student and no changes were made. U. Boardman's Winter Benchmark Improvement Report showed no significant improvement and a recommendation to change the instructional program was prescribed. The instructional change(s) resulted in much greater improvement for U. Boardman by the Spring Benchmark.

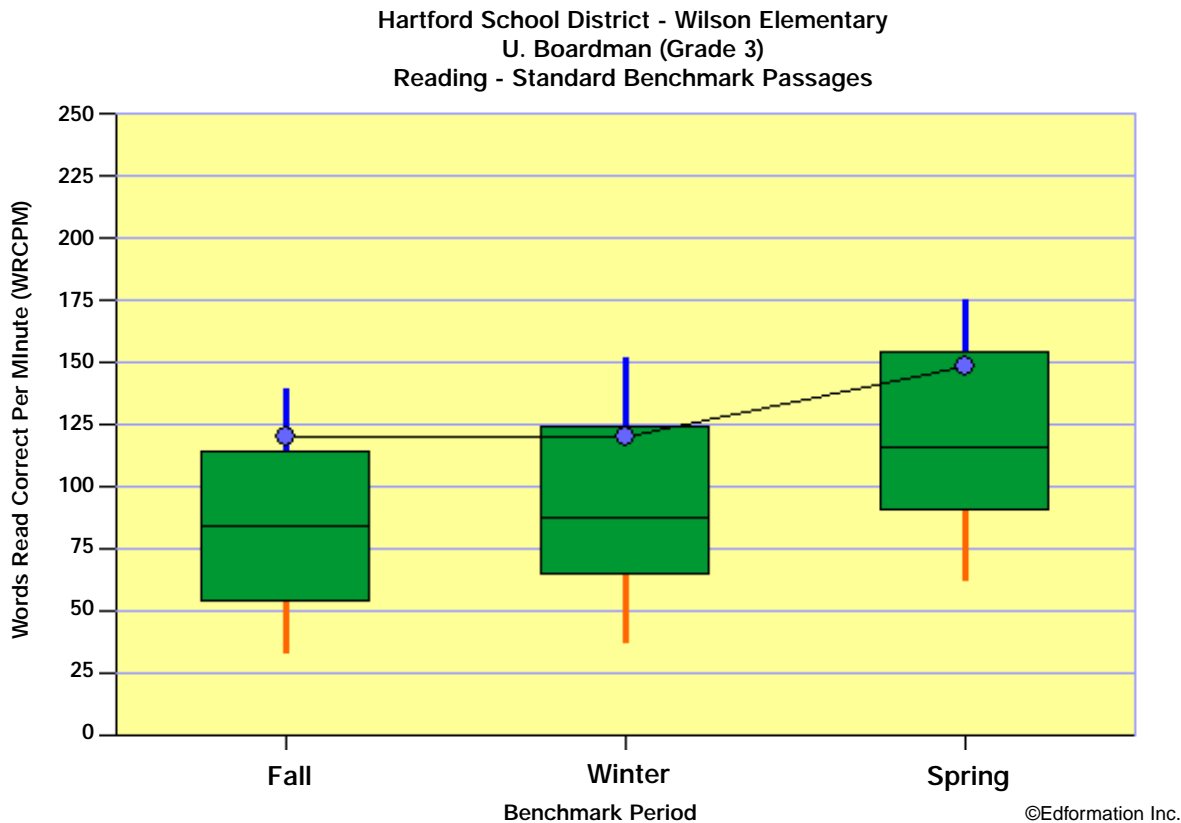


Figure 9

Qualitative Features:	Teacher Comments:
<ol style="list-style-type: none"> Has an effective strategy for unknown words. Reads more accurately. Reads "fluently" or efficiently. Reads with expression. Self-corrects errors. 	<p>U. Boardman is a High Average student compared to other students in Wilson School. He showed significant improvement when he was presented with reading material that challenged him.</p>

Benchmark Improvement Reports also include descriptive and prescriptive information about students' rates of progress (Above Average, Average, and Below Average) to increase the accuracy of teachers' Instructional Actions. U. Boardman's Winter Benchmark Improvement Report is shown in Figure 10. At the Winter Benchmark, he was not progressing satisfactorily. The instructional prescription for his teacher was to make a change in the Program. In this instance, his teacher created skill-based reading groups and placed the student in a more challenging basal reader.

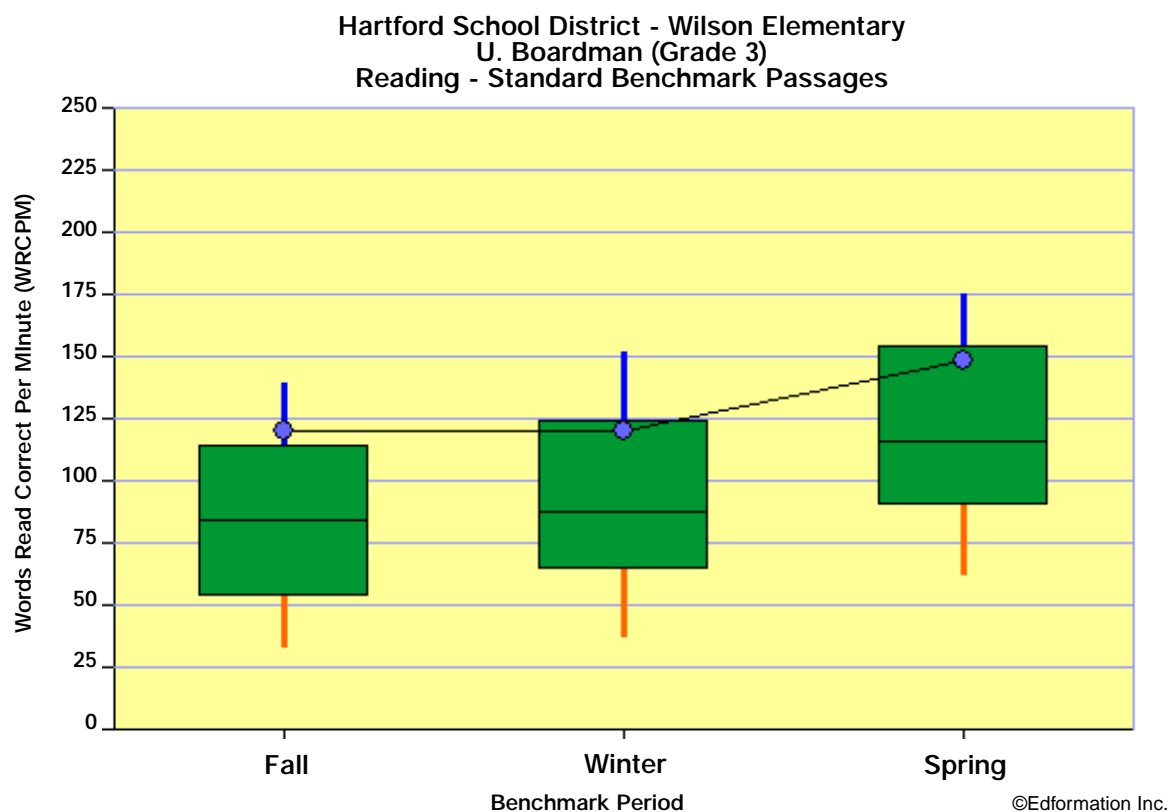


Figure 10

Qualitative Features:	Teacher Comments:
<ol style="list-style-type: none"> Has an effective strategy for unknown words. Reads more accurately. Reads "fluently" or efficiently. Reads with expression. Self-corrects errors. 	<p>U. Boardman is a High Average student compared to other students in Wilson School. He showed significant improvement when he was presented with reading material that challenged him.</p>

Formative Evaluation - Inform Instruction

Schools across the country have been collecting Benchmark information using General Outcome Measures such as Curriculum-Based Measurement (CBM) for almost 25 years. Every minute of assessment time that does not “inform instruction” is a loss of a minute of instructional time. Therefore, this Workbook is based on careful observations of educators to see what works (and what doesn’t) to make Benchmark Assessment as efficient and successful as possible. Benchmark Assessment is successful when there is a thoughtful plan that addresses 6 major topics. Within each topic, we will lay out what has been demonstrated to work and some things that have not worked well.

1. Setting It All Up
2. When to Test
3. What to Test and Score
4. Who is Tested
5. Who Does the Testing
6. How Students are Tested

Setting It All Up

Too often, assessment practices are “dropped” on teachers “from above,” with no notice, discussion, or even from the perspective of how the process helps teachers. Arguably, most of these assessment activities are only grossly intended to help teachers, and teachers are aware that students will be asked to spend a lot of time doing something they haven’t been trained well to do (or understand), that the results will be slow in coming to them, and that likely, no matter how well they have taught, the information will be used to paint a bleak picture of current education. A similar process occurs with parents. Without notification or preparation, parents often receive testing results after the testing occurs. It is little wonder, then, that as a result of poor set up, formative evaluation assessment systems designed to assist teachers, parents, and administrators make decisions proactively, efficiently, and accurately to improve instruction are painted with the same brush and often are not welcomed with enthusiasm. We see the need, then, to:

- Prepare the teaching staff.
- Prepare parents.

Preparing the Teaching Staff

We have found that the best way to deal with the problem of teachers feeling overwhelmed by testing and other external demands is to view the world from their perspective and start a set up process before rapid movement into Benchmark Assessment.

This “set up” includes providing teachers with examples of the general outcome measures and the use of Benchmark test information. For example, when educators only hear about the use of a standard oral reading fluency test as a general outcome measure of reading, we will often hear personal perspectives of the inadequacy of this type of assessment. When educators (1) see video footage of actual students reading aloud, (2) they actually assess students themselves, and (3) discuss the wealth of technical and qualitative information, this issue is usually resolved. Similarly, when educators can see actual examples of the assessment information translated into (1) useful, descriptive and prescriptive reports, and (2) a range of the instructional actions that can be taken, buy-in is enhanced.

Set up should also include a vivid and explicit focus on the “big ideas” of general outcome measurement to bring an ongoing familiarity of what is being done and why.

- Use of short, simple, time efficient “vital signs” of student achievement;
- Proactive and ongoing assessment of student growth; think about Benchmarking like physicians’ well checks;”
- Early identification of at risk students;
- Information to inform teachers about what is working and with whom it is working; and
- Benchmarking is a shared school commitment to improving outcomes for all students and communicating with parents.

These big ideas can be reinforced with key background readings such as Deno’s *Nature and Development of Curriculum-Based Measurement* (1991) or Shinn, Shinn, Hamilton, and Clark’s book chapter on using CBM in *General Education Classrooms* (2002). We have seen these big ideas posted prominently in school hallways and teachers’ lounges.

Successful set ups also communicate three “systems-change” big ideas to teachers. The first is the awareness that Benchmark Assessment may be an “add-on” to a currently cumbersome assessment system, but that one of the goals is to try to collect this kind of information with the intent to supplant other more cumbersome and less instructionally useful testing.

The second is the idea that Benchmark Assessment is a shared commitment; each teacher is not expected to implement and manage Benchmark Assessment alone. All certified personnel and support staff typically contribute their time as well. The third idea is that sufficient support will be provided to do things well; in other words, there are multiple systems to support implementation of Benchmark Assessment with integrity.

Preparing Parents

We have found it useful to also be proactive with parents using a variety of ways to communicate about Benchmark Assessment. Among the successful strategies have been (1) a letter sent directly home, and/or (2) an article in the first edition of the school newsletter describing Benchmark Assessment. Regardless of the communication approach, communication is increased by including samples of the Parent Report. An example of a letter that can be sent home from the building principal is presented in the Appendix.

When to Test

The current version of **AIMSweb Benchmark** is based on assessing students three times during the standard academic year,

Fall – during the weeks of 9/1 to 10/1

Winter – during the weeks of 1/1 to 2/1

Spring – during the weeks of 5/1 to 6/1

The time between each Benchmark Assessment is about equal, although it need not be perfectly so. The Fall Benchmark timeframe was selected based on the experiences of an initial beginning of the year

student and teacher adjustment period. Benchmark Assessment should not proceed so soon that students are not yet in the mind set of schooling, but not so late that teachers' instructional planning is delayed unnecessarily. The Spring Benchmark period was selected so that there would still be sufficient time for instructional planning and communication about instructional progress and needs to parents for the upcoming school year as well as to diminish the too often observed "end of the year" effect. The Winter Benchmark period, then was specified to allow an opportunity for a "mid-course correction" should instructional programs need changes for individual students.

It is useful to compare the suggested periods for Benchmarking to your school calendar and as a staff, discuss the actual weeks that students will be assessed. By engaging in this exercise, potential barriers (field trips, teacher in-service days, conferences) can be identified and worked around. Additionally, enhancers (getting information ready in time to share with parents at conferences, days when itinerant support staff are in the building) can be identified that make some days better than others.

Suggested Benchmark Weeks	Proposed School Calendar Week	Enhancers and Barriers
Fall (9/1–10/15)		
Winter (1/1 – 2/1)		
Spring (5/1 – 6/1)		

When the Benchmark period is identified, it is important to keep data collection confined to an interval of no more than 2 weeks from beginning to end. This timeframe is very liberal and successful implementers of an **AIMSweb Benchmark** finish in significantly less than 2 weeks.

It is expected that future versions of **AIMSweb Benchmark** will be adapted for use in those schools who use a year-round schooling model or whose grading system is set up on a quarterly system for which a parallel Benchmark Assessment timing is desirable. It is worth noting that for those cases where an increase in the number of Benchmark Assessments is needed, the **AIMSweb Progress Monitor** can be used for this purpose.

What to Test and Score

AIMSweb Benchmark currently is designed for use with the general outcome measures of oral reading (Reading Curriculum-Based Measurement; R-CBM) reading maze, spelling, Early Literacy (**AIMSweb** measures of Phonological Awareness and Phonics; DIBELS) and spelling. Mathematics computation and written expression measures are in preparation. Although there are many basic skills areas that can be assessed and used with **AIMSweb Benchmark**, we recommend that schools begin with reading (and early literacy for Grade K and beginning first grade in their first year of implementation).

By focusing on reading and early literacy, schools can work on increasing efficiency in assessment practices and using formative information to inform instruction, communicate with parents, and allocate resources in reading academic area. It is easy to add additional basic skill areas to Benchmark Assessment after "all the bugs" have been worked out in 1 area and schools are practiced in using the results.

In Reading, students are assessed individually using the standardized procedures taught in the **AIMSweb Administration and Scoring of Reading Curriculum-Based Measurement (R-CBM) for**

Use in General Outcome Measurement workbook. Students read 3 **Benchmark Standard Reading Assessment Passages**, which were carefully selected to represent graded narrative material of about equal difficulty. For more information, see **Edformation's Standard Reading Assessment Passages Technical Manual**. Although the passages were carefully constructed to minimize differences in readability, students read 3 passages so that the highest score and the lowest score, presumably from the easiest and most difficult passages respectively, are deleted. This process leaves the most representative score of the student's current reading performance, the median, as the score that is used to make decisions. For planning purposes, approximately 5 minutes should be allocated per student for (1) reading the directions, (2) doing the assessing, and (3) scoring and recording the results. This 5-minute estimate does not include the time it takes students to get to and from the testing place. This time estimate is based on the assumption that the examiner has been well trained and has practiced.

Who to Test

One important question asked frequently over the years is "Who should (and should not) be included in Benchmark Assessment?" The general principle to be used to answer this question is to assess all students on a general education teacher's class roster. This roster typically includes students who also may be eligible for and receive other educational programs such as Title I, special education, talented and gifted, and English Language Learning. These students may not receive their instruction in a given academic area (such as reading) from their general education teacher; however, it is this general education teacher who is largely responsible for (1) the student's overall general education program, especially for initiating the need for more specialized programs, and (2) communicating progress to parents.

The progress of some of these students in additional instructional programs also may be assessed as part of **AIMSweb Progress Monitor**. However, because their annual goals are more individualized and may be not be based on the rate of progress of students at their grade placement, it is important to have the independent index of their general growth provided by **AIMSweb Benchmark**.

Consistent with any assessment practice, students excluded would be pupils lacking the physical skills to participate meaningfully. For example, it would not be appropriate to have students with severe vision impairments read aloud for 1 minute.

Who Does the Testing

It was presented earlier in Setting It All Up that success in implementing Benchmark Assessment is enhanced if "the commitment is shared by the school." No where is this commitment more evident than in the identification of who does the assessment. Because the general outcome measures used as part of **AIMSweb** are time efficient, it is realistic to have general education teachers collect all the Benchmark information on each of their students. The annual commitment for assessment, not including scoring, is less than 15 minutes per year per student.

In reading, where students are assessed individually, a classroom of 25 students requires an annual time commitment of about 6 1/2 hours (25 students @ 15 minutes per year = 375 minutes). Furthermore, it is desirable for general education teachers to observe systematically how well their students perform for diagnostic purposes.

However, it is not necessary for general education teachers alone to collect all the Benchmark information. Although the total amount of time for assessment is not significant, given the quality of informa-

tion that is obtained, overall efficiency is increased by having others assist. Furthermore, general education teachers observe how their students perform, often daily, and in a variety of circumstances such that the “loss” of their own systematic observations may be offset by the gains in economy.

We recommend, then, that emphasis be placed on the economy and efficiency of assessment. It is economical and efficient to have general education teachers do Benchmark Assessment when their class can be assessed as a whole such as in spelling. In reading, it is more economical and efficient when as many trained persons as possible do the Benchmark Assessment. Within any given school, there is a range of licensed persons who can do the reading assessment, including:

Teachers in other instructional programs:

- Title I
- Special Education
- English Language Learning
- Music
- Physical Education

Service providers such as:

- School Psychologists
- School Counselors
- Speech and Language Clinicians
- School Principals/Administrators

Within any given school, there also is a range of support persons, including classroom aides, and aides for other instructional programs.

We believe that including as many of these persons communicates the shared commitment that reading and Benchmark Assessment is important not only to general education teachers, but also to the students and their parents. Imagine a general education teacher seeing the building school psychologist assessing the reading of some of the students in his/her classroom. Imagine the student who gets the opportunity to read with the building principal!

Within any given community, there is a range of persons who also may be qualified to do Benchmark reading assessment, including, but not limited to:

- Student teachers in the building or district.
- Undergraduate or graduate students in pre-service training from local colleges and universities.
- School volunteers, senior citizens, or high school students doing community service projects.

How Benchmark Testers are Trained is More Important than Who They Are

Articles in the professional literature have demonstrated that all of the groups of people can collect Benchmark information efficiently and accurately. A professional license is nice but not necessary. More important to accuracy and efficiency in Benchmark Assessment is attention to 3 variables:

1. Professional Ethics: Test Security, Confidentiality, and Sensitivity.
2. Sufficient Training.
3. Sufficient Practice and Feedback.

Professional Ethics: Test Security, Confidentiality, and Sensitivity

No assessment system can be implemented successfully without careful attention to professional ethics. We believe that it is critical that as part of training, all parties review and discuss the basics of professional testing ethics and provide the appropriate safeguards so that quality decisions are made about student performance and student learning.

Foremost among the professional ethics issues is test security. Educators must do what they can to ensure that the tests that are used with students measure what students have learned, not just how they have been prepared for testing. We have all witnessed concern about “teaching to the test,” an all too uncommon reaction to the pressures of high-stakes accountability systems. In our experience, however, we have witnessed time-and-again the professional integrity of educators in using general outcome measures for the purposes for which they were intended and few experiences of teaching to the general outcome measures.

That said, we believe it important to spend training time reviewing these issues. The content of the specific probes is to be protected and all reasonable attempts should be made to enable them to function as measures of student learning. In other words, all professional safeguards should be made to prevent teaching to the specific probes.

It is an appropriate option to:

1. Record how students performed on specific Benchmark probes for purposes of analysis by teachers for follow up instructional planning purposes.
2. Maintain Benchmark probes in confidential student achievement records for purposes of comparing student growth over time.

Test security is violated if:

1. Persons provide direct instruction or opportunities to practice the content of the Benchmark Assessment Materials.
2. Parents are provided copies of their children’s Benchmark protocols unless they are based on the legal standards for parental rights to information as part of the local school, state or federal practices.

A second important issue is one of confidentiality and sensitivity. That is, comments about how individual students performed, either positive or negative, are to be made only under professional circumstances, when the information can be used to inform teaching. For licensed professionals, simply providing a set of positive and negative examples as part of training is sufficient. For example, a school psychologist who tests some second-grade students may comment generally about what good readers they are. If referring to Miguel, any comments should be made first privately to Miguel’s teacher such as:

“I noticed that Miguel was a great reader and really used word attack strategies well,” or

“I observed that Miguel seemed to lose his place a lot and therefore the scores I obtained may not be accurate.”

It is not professional to discuss Miguel’s performance casually in the lunchroom.

Another professional place is in meetings planned for discussing students’ educational needs.

For example, at an annual IEP review meeting, a special education teacher may comment that when they assessed Sandra as part of Benchmark Assessment,

“She was now reading almost as well as most students in her classroom.”

For non-licensed personnel, like aides, teachers in training, or volunteers, at least 10 minutes of training and discussion should be devoted to the topic. For these persons, it is not appropriate to discuss the positive or negative performance of any individual child except with the classroom teacher. It is important to emphasize that what seems as a casual conversation can have serious consequences. Thus going over situations like...

You saw your neighbor in the grocery store and he asked you how the Benchmark Assessment went and *“Did you assess my son?”*

Answer?

“As part of what I did, I shouldn’t speak about individual children, including whether I assessed them.”

Alternately, in supervision meetings, student teachers also should not discuss the performance of individual students unless (1) they use a false name, and (2) they have discussed the information with their supervising teacher.

Training

Regardless of licensure or prior assessment experience, persons who will be doing Benchmark Assessment should be well trained. We suggest allocating approximately 10 - 15 minutes at the beginning of training to review issues of Professional Ethics, with practical examples.

The specific content and scoring is laid out in each of the general outcome measure’s training workbook. Each trainee should have their own workbook(s) as it includes practice examples that must be completed. As some general guidelines for training, we suggest the following:

- 1. It is better to do too much training than too little training.** Although the Benchmark Assessment procedures are generally simple to do, they are not automatically easy to learn. Even when initial training is intensive, we expect that there will be errors in implementation that must be corrected through feedback.
- 2. The training atmosphere should be one where everyone provides, and is open to, feedback.** This means systematically asking trainees to provide their scores on the practice exercises. For example, if trainees are sitting 5 to a table, 2 tables can give their scores for Reading Practice Exercise 1, 2 other tables for Reading Practice Exercise 2, etc.
- 3. Everyone needs opportunities for corrective feedback.** Trainees should pair up and complete one *Accuracy of Implementation Rating Scale (AIRS)* on each other. After feedback is provided, a second practice administration should occur.
- 4. Complete the exercises for Inter-Scorer Agreement and post the scores.** Doing so, again, will provide opportunities for corrective feedback and gaining confidence in the quality of the Benchmark information.

Practice

If the recommended training materials and exercises are completed as detailed, trainees will be accurate in their administration and scoring. However, without practice, on average, trainees will be slow and inefficient. In reading, studies completed by Dr. Caren Wesson documented that experienced teachers with good training in administration and scoring averaged about 5-7 minutes to read the directions, have the students read for 1 minute, score the results, and write down the score. With practice and peer feedback, teachers increased their assessment efficiency to an average of about 2 minutes total for a 1-minute reading sample. We suggest that trainees work with a peer and jointly administer and score 8 - 10 reading passages. Again, the AIRS can be used to provide structured feedback.

How Students Are Tested

As with most human endeavors, it is the quality of planning that ensures success. Efficiency of Benchmark Assessment also is the result of quality planning. We see 4 major tasks that must be accomplished.

1. Appointing a Benchmark Assessment Coordinator
2. Creating an Assessment Schedule
3. Preparing the Assessment Materials
4. Setting Up Places to Assess

Appointing a Benchmark Assessment Coordinator

One person in each school should be appointed the Benchmark Assessment Coordinator. Although Benchmark Assessment may be part of a school district initiative where the overall process, including assessment training, is managed outside the school, having 1 building representative work with school staff is still desirable. The Coordinator can manage the planning process using a ***Benchmark Assessment Essentials Checklist*** such as the one illustrated on the following page.

Benchmark Assessment Essentials Checklist:

Task	Person Responsible	X
Staff Prepared		
1. Introductory Presentation		
2. Background Reading Materials		
Parents Prepared		
1. Introductory Letter/Article		
2. Open House Presentation		
Benchmark Assessment Periods Identified and Marked on School Calendar		
Persons Doing Assessment Identified and Contacted		
1. Training Session Planned		
2. Training Materials Prepared		
3. Training Completed		
4. AIRS Completed		
5. Inter-Scorer Agreement Completed		
6. Assessment Practice Completed		
Classroom Rosters Obtained and Student Participants Identified		
Benchmark Assessment Materials Prepared		
1. Student Materials (e.g., unnumbered Reading Assessment Passages)		
2. Examiner Materials		
Assessment Directions		
Examiner Materials (e.g., numbered Reading Assessment Passages)		
Stopwatches		
Scoring Keys		
Assessment Sites Identified		
Assessment Schedule Prepared		

Benchmark Assessment Essentials Checklist (continued)

Task	Person Responsible	X
Assessment Schedule Distributed to All Testers		
Trouble-Shooter Identified		
Process of Collecting Tests Identified and Disseminated		
Process for Entering Student Information Identified		
Student Information Entered		
Process for Supporting Teachers in Obtaining Results Identified		
Process for Communicating Results to Parents Identified		
Benchmark Assessment Debriefing Planned		

Preparing a Printed Testing Schedule

A printed testing schedule that is disseminated to all testers and posted prominently in a place that can serve as a Benchmark Testing Center (e.g., staff lunchroom) makes it clear which testers are assessment which students, where, and when. A simple format such as the one below can be organized: See Appendix.

Sample Reading Benchmark Testing Schedule
Thursday, Jan. 19th

Grade/Teacher/Classroom/ # Students	Time	Testers	Location
Grade 1 Scalia-Room 107 (22) Lopez-Room 108 (20) Chang-Room 109 (23)	9:00–9:40 9:50–10:30 10:40–11:20	Scalia, Lopez, Anderson, Chang	A-Wing Testing Center
Grade 2 Rhenquist-Room 116 (28) Dailey-Room 117 (28) Marcus-Room 118 (25)	9:00–9:40 9:50–10:30 10:40–11:20	Rhenquist, Marcus, Antonelli, Goodale, Parker Dailey, Marcus, Antonelli, Goodale, Parker Dailey, Marcus, Antonelli, Goodale, Parker	B-Wing Testing Center B-Wing Testing Center B-Wing Testing Center
Grade 3 Bush-Room 126 (29) Cheney-Room 127 (24) Gore-Room 128 (25)	9:00–9:40 9:50–10:30 10:40–11:20	Bush, Petitit, Thurber, Peinado, Piper Cheney, Petitit, Thurber, Peinado, Piper Gore, Petitit, Thurber, Peinado, Piper	Media Center Media Center Media Center

In this example, which shows the testing schedule for Grades 1-3 at the Winter Benchmark, there are three classrooms per grade. Reading tests are conducted by classroom teachers and 3-4 members of the school staff, including support staff and trained paraprofessionals. All three grades are tested simultaneously and all tests will be finished in one morning. Approximately 40 minutes is allocated for testing each classroom, based on an estimation of about 5 minutes per student. In Mr. Marcus' classroom of 25 students, five people will be testing. Each examiner will test about five students, leaving about 15 minutes of additional time to allow students to come to and leave the reading station, and to give examiners some flexibility. Each examiner has the names of the students they will be testing written on an examiner copy for each student. Testing will be conducted either at small desks in the quiet hallway outside the classroom or in the media center. After each student is tested, the student will be given the name of the student that will be tested next, and will direct that student to the reading station. Grades 4-6 would be scheduled for testing in a similar manner the next morning.

Preparing the Testing Materials

After the specific materials are identified, the testing materials are prepared.

For student testing materials in the area of reading and/or early literacy, we suggest that their unnumbered copies be laminated because they will be handled over and over by many children throughout the year. Another suggestion that saves considerable time is to print each grade's testing materials on a different, light color paper. Using lightly colored student passages will not affect their reading scores and helps keep the testing materials organized for the examiners.

For examiners, we suggest preparation of three sets of materials. The first set is comprised of a set of standardized directions. Again, we suggest laminating these directions because they are handled by many examiners during the academic year.

The second is a copy of the class rosters. These are obtained from **AIMSweb Benchmark**. See Appendix for example of Benchmark Assessment Sheet. The third set of examiner materials consists of the numbered **Standard Benchmark Reading Assessment Passages** for each grade. These materials can be color-coded the same as the student copies. The color-coding helps the examiners keep their materials organized and assists in the organization and entry of the assessment results. One option for organizing the testing materials is to prepare individual student "portfolios" that include the examiner copies of the 3 Benchmark Reading Assessment Passages on an 8-1/2" x 11" piece of paper folded in half lengthwise. Typically the type size of the examiner copy is reduced to fit the page.

Setting Up Places to Test

Setting up places to conduct Benchmark Assessment in Reading is straightforward. When group testing, as in spelling, it is best to use students' own classrooms. No time is wasted moving students to places to be tested. When reading is assessed, more careful consideration is required. Although the reading testing environments are flexible, two goals should be attained. First, the environment should be reasonably quiet and away from distractions (sinks, water fountains) and include a table or desk so that the examiner can sit across from, not next to, the student and so that the student can have a place to lay down their copies of the reading passages. Second, all attempts to minimize the distance students must travel should be made. For example, a media center is a good place to test students. It usually is quiet and has a number of tables that are appropriately spaced for more privacy. However, if examiners must wait for students who were tested to return to their classrooms to get the next students to be assessed, valuable professional time is wasted. If the media center is not close to the classroom, it may be best to bring the whole class to the media center and allow them to read their own books privately while waiting to go to reading stations for testing. The same issues arise when using a cafeteria, gym, or empty classrooms. Examiner and student time should be spent testing and reading, not waiting and walking.

Reading stations in the hallway outside of classrooms are suitable providing scheduling takes into account when there may be student traffic is suitable. Finally, when there are not a large number of examiners, two to three reading stations can be created in classrooms that provide the necessary quiet and privacy for reading.

After Testing: What Then?

After the Benchmark Assessment information is collected, we see the need for 3 major activities:

- Assuring Information is Entered Quickly and Accurately as Possible.
- Providing a Support System to Help Professionals Use Information.
- Complete a Benchmark Assessment Debriefing to Improve Efficiency.

To ensure that the Benchmark information can be used as intended, there should be a plan to enter the information by teacher by grade as quickly as possible. Specific procedures for entering the information are described in **AIMSweb Benchmark** documentation. With **AIMSweb Benchmark**, it is also possible to print Benchmark Assessment scoring sheets at the grade, class, and teacher levels.

However, the specific data entry approach should be clear to all and deadlines established. In some instances, classroom teachers enter their students' Benchmark information. In other instances, a designated person(s) enters all the information. These decisions should be made before, not after Benchmark Assessment.

As we stated earlier, we believe that it is important to support the users of the Benchmark Assessment information; that is, we should not just provide Benchmark Improvement Reports but schools should have a process to help professionals *use* the information. This process can be initiated by having a post-Benchmark Assessment review of the information and having select teachers present for discussion some of the reports that were generated. Opportunities for individual consultation sessions should be considered as well.

Finally, we have observed that Benchmark Assessment becomes considerably more efficient with practice. Benchmark Assessment is a time efficient way to collect information useful for informing teaching, communicating with parents, accountability, and resource allocation. However, the time spent in the first Benchmark Assessment is typically much longer than that spent in subsequent Benchmarking. We have witnessed the amount of test time decrease by 50% from the first Benchmark to the second Benchmark as a result of gaining some experience and having the opportunity to "practice" on a large scale. Efficiency is increased when school staff is given the opportunity to go over what worked, what didn't work, and what could be done to increase efficiency. We see this topic worthy of discussion as part of a post-Benchmark Assessment staff meeting.

Summary

Benchmark Assessment using simple, but important general outcome measurement provides a data base to teachers, administrators, and parents for making decisions about the growth and development of basic skills. The development of Benchmark Reports at the teacher level helps inform instruction so that students at risk can be identified early and programs adapted to meet their instructional needs. Benchmark Reports for Teachers can help plan instruction and as importantly, help evaluate "what works" for given students. Benchmark Improvement Reports at the parent level can increase home and school communication and collaboration, providing simple to understand information about student performance and progress. Benchmark Reports for Principals and Administrators can help in the difficult tasks of accountability and resource allocation.

This workbook has provided the underpinnings of Benchmark Assessment, from how to "set up" to planning data collection, to determining what to do after Benchmark Assessment is completed. Strategies that have been successful in schools around the country were described to enhance the success of your Benchmark Assessment experiences.

References

- Deno, S. L. (1992). *The nature and development of Curriculum-based Measurement*. *Preventing School Failure*, 36(2), 5-10.
- Shinn, M. R. (1989). *Curriculum-Based Measurement: Assessing Special Children*. New York, NY: Guilford Press.
- Shinn, M. R. (1998). *Advanced Applications of Curriculum-Based Measurement*. New York, NY: Guilford Press.
- Shinn, M.R., Deno, S.L., & Fuchs, L.S. (2002). *Using Curriculum-Based Measurement in a Problem-Solving Model*. New York: Guilford.
- Shinn, M.R., Shinn, M.M., Hamilton, C., & Clark, B. (2002). *Using curriculum-based measurement in general education classrooms to promote reading success*. M.R. Shinn, H.M. Walker, and G. Stoner, *Interventions for Achievement and Behavior Problems II: Preventive and Remedial Approaches*. Bethesda, MD: National Association of School Psychologists.
- Wesson, C. (1987). *Facilitating the efficiency of ongoing curriculum-based measurement*. *Teacher Education and Special Education*, 9, 166-172.

Books Available at www.guilford.com or www.amazon.com

Appendix: A Sample Letter to Parents

A sample letter to parents explaining a Benchmark Assessment system, similar to the letter below, is also available for printing through AIMSweb Benchmark.

Dear Parent(s):

As our children grow, most of us see charts like the one to the right. Our doctors weigh our children and measure how tall they are. This information is put on a chart like the one to the right. This chart shows where our child stands compared to other children and how fast they are growing. These charts help doctors to understand if our children are growing properly or if something may be wrong, enabling them to do something to help as soon as possible.

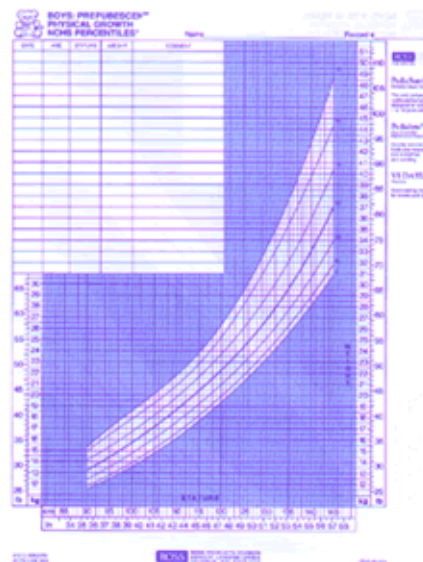
Doctors use children's height and weight as "Well Checks", or vital signs of development. We all know how important it is for our children to "grow" in reading as well. In fact, successfully learning to read at an early age is probably one of the most important factors in a child's learning, both inside and outside of school.

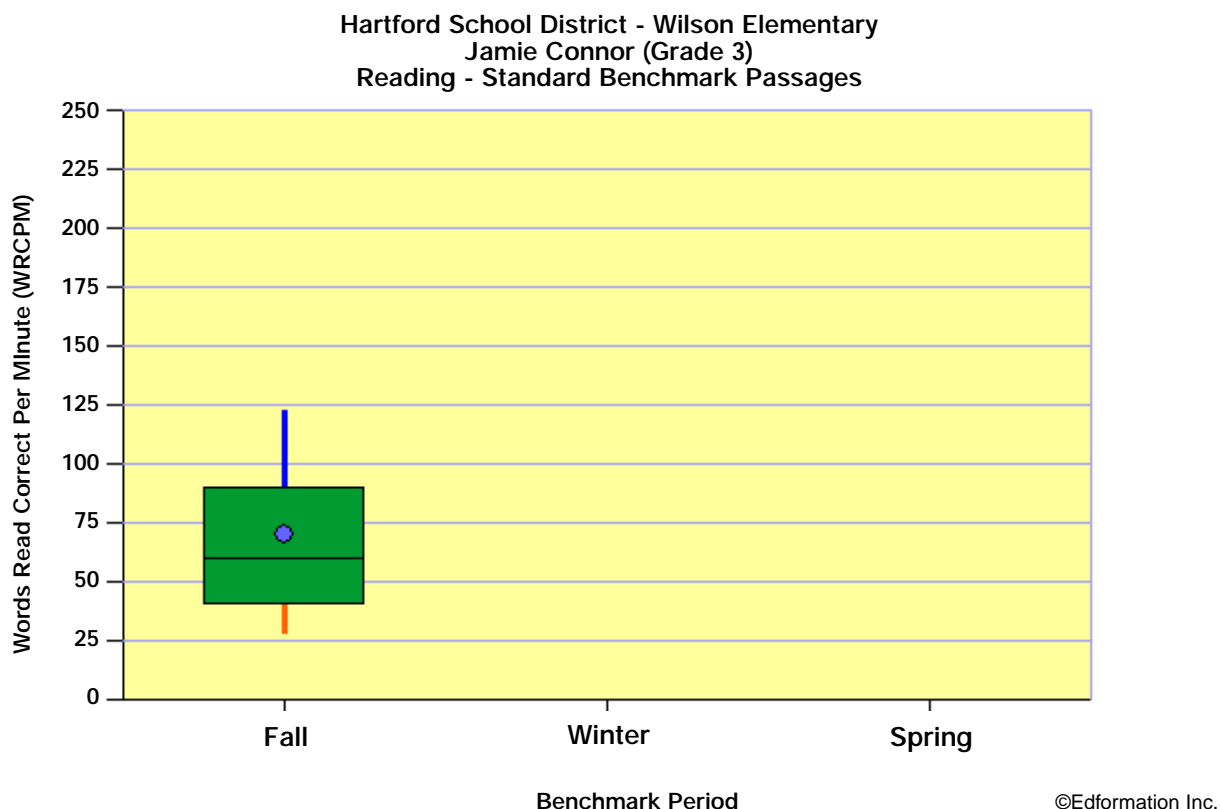
Because reading and reading growth are so important, on a regular basis, our school is going to give you "Well Check" information indicating how well your child is reading. We now know what the goals of the reading "Well Checks" should be, and will give this information to you three times per year. Generally, the first "Well Check" is given within the first few weeks of school.

The reading "Well Checks" are simple and accurate. We will have your child read 3 stories from passages comparable to their grade level and we will count the words they read correctly, as well as how well they read the stories. The "Well Checks" take less than 5 minutes. We like to think of it as taking your "child's reading temperature".

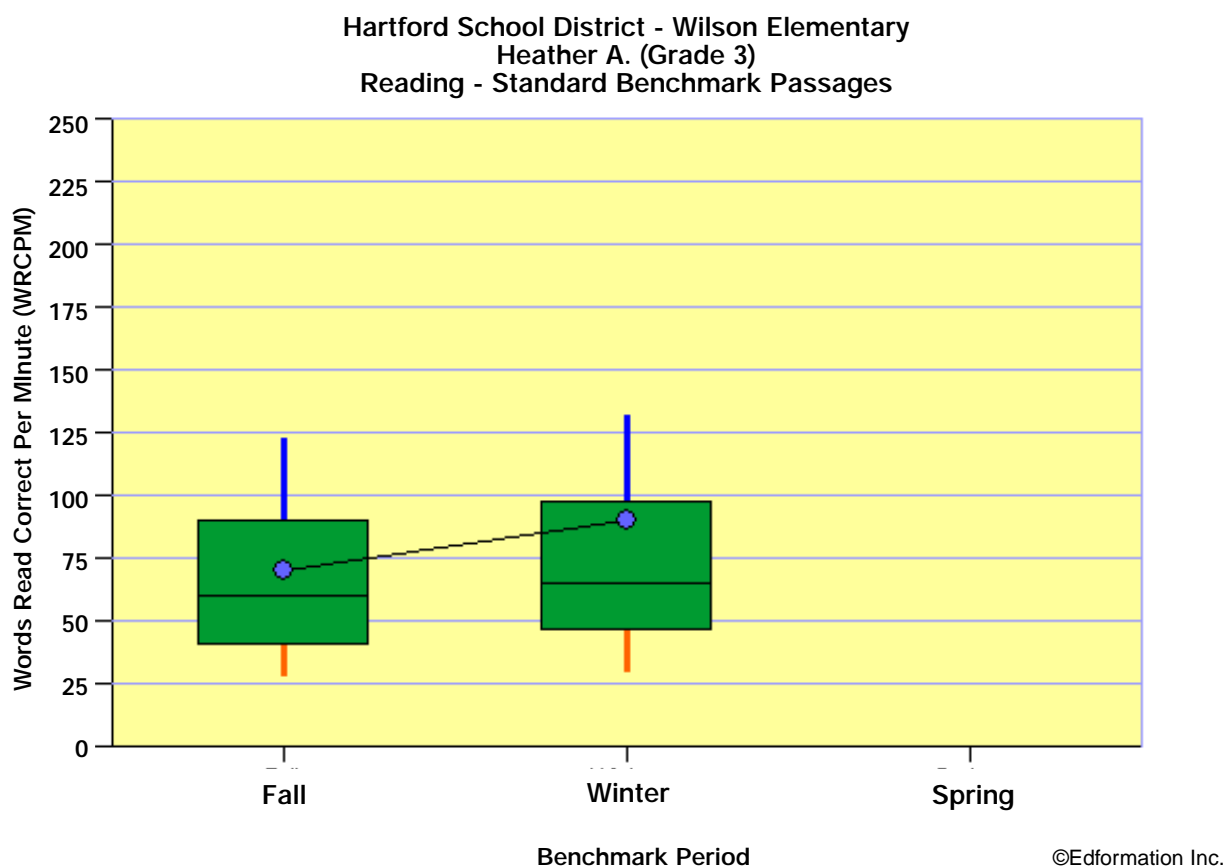
Teachers will use these "Well Checks" to help them plan their reading instruction. This information will also help them individualize for students who are above or below average in their reading abilities. The "Well Check" information for your child will be put on a chart like the one below, and your child's teacher will explain it to you at conference time.

This chart shows how one student read the stories compared to other classmates in the same grade. This chart also displays how well an average third-grader reads at our school and reads third-grade stories at the beginning of the year (Fall). The number of words this child read correctly in 1 minute is indicated by the blue dot. The black horizontal line is the score of what the very middle student read and the green box displays scores of average third-grade readers. Scores within the blue line are above average readers, while scores within the red line are below average readers. This graph indicates the student read about 70 words correctly in 1 minute. The average reader in this grade read about 65 words correct per minute.





In January (Winter), your child's teacher will do another "Well Check" to see if an improvement in your child's reading ability has been made. Another chart, similar to the one below, will then be given to you. This chart will indicate what your child's score was in the fall and what your child's score is in Winter. The reading "Well Check" chart below shows the rate of progress of an average student in the third grade. The lines connecting the blue dots show the rate of progress. Notice an improvement from 70 words correct in Fall to about 90 words correct in Winter. This reader is growing at a slightly faster rate than other students in our school. This information shows success!



These reading "Well Checks" will be done again during the last several weeks of school (Spring) and the results will be shown to you. We will then be able to inform you of how much reading growth your child has made throughout the year, and what may be needed in order for your child to become a truly successful reader.

As a school staff, we are excited about our ability to know where all of our students are at on their path to being successful readers and learners. We hope that you are excited as well. We look forward to seeing and talking with you regarding the progress of your child.

If you have any questions about our Reading Improvement Report Card and Reading Well Check Program, please feel free to contact me.

Reading Benchmark Testing Schedule

Grade/Teacher/Classroom/ # Students	Time	Testers	Location
Grade K			
Grade 1			
Grade 2			
Grade 3			
Grade 4			
Grade 5			
Grade 6			

Benchmark Assessment Essentials Checklist

Task	Person Responsible	X
Staff Prepared		
1. Introductory Presentation		
2. Background Reading Materials		
Parents Prepared		
1. Introductory Letter/Article		
2. Open House Presentation		
Benchmark Assessment Periods Identified and Marked on School Calendar		
Persons Doing Assessment Identified and Contacted		
1. Training Session Planned		
2. Training Materials Prepared		
3. Training Completed		
4. AIRS Completed		
5. Inter-Scorer Agreement Completed		
6. Assessment Practice Completed		
Classroom Rosters Obtained and Student Participants Identified		
Benchmark Assessment Materials Prepared		
1. Student Materials (e.g., unnumbered Reading Assessment Passages)		
2. Examiner Materials		
Assessment Directions		
Examiner Materials (e.g., numbered Reading Assessment Passages)		
Stopwatches		
Scoring Keys		
Assessment Sites Identified		
Assessment Schedule Prepared		

Benchmark Assessment Essentials Checklist (Continued)

Task	Person Responsible	X
Assessment Schedule Distributed to All Testers		
Trouble-Shooter Identified		
Process of Collecting Tests Identified and Disseminated		
Process for Entering Student Information Identified		
Student Information Entered		
Process for Supporting Teachers in Obtaining Results Identified		
Process for Communicating Results to Parents Identified		
Benchmark Assessment Debriefing Planned		